21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

Nonlinear Impulsive Systems: A Causal Model*

Erik Verriest[1]

***Abstract*— Systems driven by impulsive inputs are readily modeled for linear systems. Extension of the results to nonlinear systems has run into many problems. Consequently impulsive nonlinear systems are usually described by the *effect* of impulses, thus giving rise to jumps in the state space. However, it is not clear a priori how these jumps are to be related to the effective impulsive inputs (the *causes*) to the system. We derive such results for bilinear systems of arbitrary order, and discuss further extensions to other classes of systems. The method is justifiable by taking an approach touching on non-standard analysis. It is based on *first principles* considering singular functions as sequences of regular functions, pretty much as they are thought in an undergraduate signal course, and leads to a definition of *insensible* times and functions. The latter provide the *fine structure* extension.**

I. INTRODUCTION

Impulsive inputs are very convenient as approximations for *fast* inputs, in linear systems. Here fast means of short duration compared to the natural time constants (or periods) in the system. These models are of interest in biomedical and ecological applications [3], [5], [27], [30], and in aerospace engineering systems [15], [18]. Indeed, in such cases the error introduced by replacing the actual input by an impulsive one is small, but the computational burden in computing the response is greatly simplified. For this reason it is extremely desirable to extend this computation with impulsive inputs to the nonlinear realm. Unfortunately, many such attempts have failed, mainly because of the fact that the Schwartz distribution theory does not allow multiplication of distributions. Whereas there would not be any problem with $\delta(t-t_1)\delta(t-t_2)$ when $t_1 \neq t_2$, it is the zero function, products like $\delta^2(t)$ cannot be defined. Several other peculiar things happen: for instance the Heaviside, or unit step function, $H(t)$, is invariant in the class of piecewise continuous functions if it is raised to some power. However, application of the chain rule leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}H^n(t) = nH^{n-1}(t)\delta(t),$$

which then implies in turn that $nH^{n-1}\delta(t)$ should not depend on n, and in fact equate to simply $\delta(t)$. But this is inconsistent with $H^n(t) = H(t)$. Furthermore, impulsive control poses no problem for time varying linear systems. See for instance the work of Silverman [25], or Kurzhanski and Osipov [14]. It has been shown that for the smooth time varying model

$$\dot{x}(t) = A(t)x(t) + b(t)u(t),$$

the impulsive input $u(t) = u_0\delta(t) + u_1\dot{\delta}(t) + \cdots + u_{n-1}\delta^{(n-1)}(t)$, generates a jump in the state of the form

$$\Delta x = R(A(t), b(t))[u_0, \ldots, u_{n-1}]^\top,$$

where

$$R(A(t), b(t)) = [b(t), (A(t)-\mathbf{D})b(t), \ldots, (A(t)-\mathbf{D})^{n-1}b(t)]$$

is the time varying reachability matrix. The differential operator $\mathbf{D}$ is considered to act towards the right [29]. Thus

$$\begin{aligned}(A-\mathbf{D})b &= Ab - \dot{b}\\ (A-\mathbf{D})^2 b &= A^2 b - 2A\dot{b} - \dot{A}b + \ddot{b}.\end{aligned}$$

Note that this compact form in terms of powers of the operator $(A-\mathbf{D})$ is a lot more insightful that its expanded form. However, once nonlinear models are introduced, say

$$\dot{x} = f(x) + g(x)u, \qquad (1)$$

then the problem of allowing impulsive inputs becomes problematic, as essentially x and thus also $g(x)$ must contain a discontinuity at the very moment the impulse acts on the system. For this reason, nonlinear impulsive systems are typically described by the *effect* of the impulse on the state, a state dependent jump,

$$x(t_+) - x(t_-) = J(x(t_-), u_0)$$

for the function J is assumed to be given, and u_0 is a jump parameter encoding somehow the impulse induced by the input, see for instance the seminal work of Bainov and Simeonov [1], and the recent monographs [31], [32]. There is, however, no direct connection given between this J and the smooth dynamics in (1). The present paper seeks to derive such a *causal* analysis. This paper gives a reinterpretation to the work on optimal control involving measures [17], [21] and independently Orlov [19].

One way out of the impasse with regards to $\frac{\mathrm{d}}{\mathrm{d}t}H^n(t)$ is to postulate that microscopically the $H^n(t)$ are different functions for all n. This means that one should encode a rate of convergence of how fast a sequence of regular functions should approach the limit function, which brings one to the realm of nonstandard analysis [22]. Indeed, classically, a generalized function can be described as a weak limit of regular functions [11]. In this sense, given a convergent sequence $(f_1(t), f_2(t), \ldots, f_n(t), \ldots)$ of regular (continuous) functions, its weak limit is associated with the sequence, or more precisely the *equivalence class* of all weakly converging sequences with the same weak limit. This is similar with defining irrational numbers as equivalence classes of sequences of rational numbers. However, such an

[1]Erik Verriest is with the Faculty of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA
erik.verriest@ece.gatech.edu

ISBN: 978-90-367-6321-9

equivalence class can be partitioned further according to the *rate of convergence* which gives rise to the nonstandard reals. The focus in this paper will be on operator methods and is inspired by the work in [2], [23], [24], [28] and applications in electromagnetics [20]. An outline of this fine structure approach is presented in the appendix.

II. MOTIVATING EXAMPLE

Consider the simple scalar nonlinear system

$$\dot{x} = xu,$$

with initial condition x_0. We note that if $x_0 = 0$, no other states can be reached regardless the magnitude of the input. We postulate that for an impulsive input $u(t) = u_0\delta(t)$ this also holds. However, when $x_0 \neq 0$, then

$$\frac{\dot{x}}{x} = u_0\delta(t)$$

from which

$$\log x(0+) = \log x(0-) + u_0$$

and thus $x(0-) = x_0 \neq 0$, $x(0+) = \mathrm{e}^{u_0}x(0-)$, and we note that this is indeed consistent with the postulated case where $x_0 = 0$.

III. GENERALIZATION

At once we can extend this to the arbitrary scalar case, linear in the control:

$$\dot{x} = g(x)u.$$

Now, an impulsive input yields

$$\frac{\dot{x}}{g(x)} = u_0\delta.$$

Letting G be the primitive of the reciprocal g^{-1} (not the inverse function), this yields

$$G(x(0+)) - G(x(0-)) = u_0.$$

If $G(\cdot)$ is invertible, with inverse function denoted by $\overleftarrow{G}$, then we obtain the effect of the impulse in explicit form

$$x(0+) = \overleftarrow{G}(u_0 + G(x(0-)).$$

This is related to Abel's functional equation (See [12]).
Example: The logistic system: $\dot{x} = x(1-x)u$ where the input, u is proportional to the carrying capacity, has impulse effect

$$x(0+) = \frac{x(0-)\mathrm{e}^{u_0}}{1 - x(0-) + x(0-)\mathrm{e}^{u_0}}.$$

This is displayed in figure 1: Initial states $x(0-)$ are located on the horizontal axis through $u = 0$, and the corresponding $x(0+)$ can be read off as function of u (vertical axis).

Alternatively, we can approximate this ODE by its Euler discretized form for a time step ϵ:

$$x((k+1)\epsilon) = x(k\epsilon) + g(x(k\epsilon))u(k\epsilon)\epsilon \qquad (2)$$

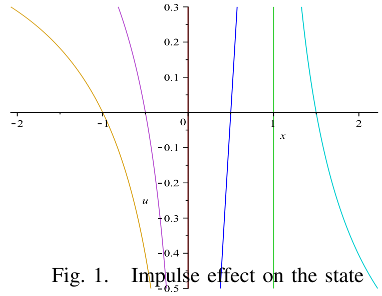

Fig. 1. Impulse effect on the state

Let an impulse of strength u_0 be approximated by a square pulse of duration $\tau = N\epsilon$, and magnitude u_0/τ, so that the "area" is conserved when taking the limit for $N \to \infty$. We then have the N-fold iterated function

$$x(N\epsilon) = f^{\circ N}(x(0)),$$

where $f(x) = x + \epsilon g(x)\frac{u_0}{\tau} = x + \frac{u_0}{N}f(x)$, and $f^{\circ k}$ is defined by

$$f^{\circ(N+1)}(y) = f(f^{\circ N}(x)) = f^{\circ N}(f(x))$$

with $f^{\circ 1}(x) = f(x)$. Letting id be the identity operator, where $\mathrm{id}(y) = y$, this can be expressed as

$$x(N\epsilon) = f^{\circ N}(x(0)) = \left(\mathrm{id} + \frac{u_0}{N}f\right)^{\circ N}[x(0)].$$

Now let $N \to \infty$ to get

$$x(\tau) = \lim_{N\to\infty}\left(\mathrm{id} + \frac{u_0}{N}f\right)^{\circ N}[x(0)].$$

This is independent of τ, so that we can now let $\tau \to 0$ to get, in a notation analogous to

$$\lim_{N\to\infty}\left(1 + \frac{a}{N}\right)^N = \mathrm{e}^a$$

the operator symbol (an exponential iteration)

$$x(0+) = \mathrm{e}^{\circ u_0 f}[x(0-)].$$

This should not be confused with the exponential series expansion of $\mathrm{e}^{u_0 f(x)}$.

Finally a third, more customary representation, is given by the Lie transport form. Indeed, if g satisfies a Lipshitz condition, then locally the initial value problem for $\dot{x} = g(x)$ can be solved. Let the solution at time t for initial value y at time 0 be denoted by $\phi(t, y)$. Thus $\phi(0, y) = y$. Let h be an arbitrary function in some space $\mathcal{H}$ and consider the induced operators

$$\Phi_t : \mathcal{H} \to \mathcal{H}, \quad (\Phi_t h)(y) = h(\phi(t, y))$$

then it is well known that the Φ_t form a semigroup with the Lie-derivative, $L_g = g\frac{\partial}{\partial x}$, as infinitesimal generator. Thus the solution of the nonlinear equation may also be represented by (see appendix)

$$h(\phi(t, y)) = \sigma_y \mathrm{e}^{tL_g}h$$

where σ_y is the *evaluation functional* at y. The effect of the impulse is then easily obtained as

$$x(0+) = \sigma_{x(0-)} \mathrm{e}^{u_0 L_g} \mathrm{id}.$$

The first terms in this series are

$$x(0+) = x(0-) + u_0 g(x(0-)) + \frac{u_0^2}{2!} g(x(0-)) g'(x(0-)) + \dots$$

This is the celebrated Magnus expansion, which also lies at the heart of the Runge-Kutta ODE solvers. See [2], [23], [24], and for a combinatorial proof with pictograms: [28].

IV. BILINEAR FIRST ORDER SYSTEMS

Consider first a system of n coupled first order ODE's, in vector form

$$\dot{x} = Axu. \tag{3}$$

Let us approximate the Dirac delta of strength u_0 by an *arbitrary* function $u(t)$ of finite duration, τ, with the single constraint that $\int_0^\tau u(t)\,\mathrm{d}t = u_0$. Integrating (3), we find

$$\mathrm{d}x = Axu\,\mathrm{d}t.$$

Let $u\,\mathrm{d}t = dU(t)$, and set $x(t) = y(U(t))$ with $U(0) = 0$, then the integral from 0 to t is

$$y(U(t)) = \mathrm{e}^{AU(t)} y(U(0)).$$

For $t = \tau$, this implies with $U(\tau) = u_0$, that

$$x(\tau) = \mathrm{e}^{Au_0} x(0).$$

Since the right hand side is independent of τ, letting $\tau \to 0$ gives the jump $x(0+) = \sigma_{x(0-)} \mathrm{e}^{Au_0} \mathrm{id}$, and this regardless the precise approximation of the impulse. This result is also corroborated by the exponential of the Lie derivative of the form $L_{Ax} g = \frac{\partial g}{\partial x} Ax$.

Consider now $F : \mathbb{R}^n \to \mathbb{R}^n$, defining the single input controlled dynamical system $\dot{x} = F(x)u$. Again, it is easily seen that by changing t to $U(t)$, and setting $x(t) = y(U(t))$, the expression

$$y(1) = \mathrm{e}^{u_0 L_F} y(0)$$

results, which means $x(0+) = \sigma_{x(0-)} \mathrm{e}^{u_0 L_F} \mathrm{id}$, and this is again independent of the precise structure of $u(\cdot)$. The independence of the above results from τ is traced back to the fact that outside $[0, \tau]$ the input is zero, so that the state remains stationary there.

For a multi-input bilinear system of the form

$$\dot{x} = \sum_{i=1}^m A_i x u_i(t)$$

where $u_i(t) = u_{i0} \delta(t - t_i)$, with $t_1 < t_2 < \cdots < t_m$, it follows then readily that

$$x(t_m+) = \mathrm{e}^{A_m u_{m0}} \cdots \mathrm{e}^{A_1 u_{10}} x(0).$$

However, if $t_j = t_{j+1}$, while all other strict inequalities are maintained, then

$$x(t_m+) = \mathrm{e}^{A_m u_{m0}} \cdots \mathrm{e}^{(A_{j+1} u_{j+1,0} + A_j u_{j0})} \cdots \mathrm{e}^{A_1 u_{10}} x(0),$$

provided that the two delta functions may be approximated by the *same* regular function u. Unless A_j and A_{j+1} commute, this will differ form the case with strict impulsive time inequalities. For the multi-input case, $\dot{x}(t) = F(x)u$ with $F : R^n \to R^{n \times m}$, uniqueness is guaranteed by a Frobenius condition. See [17], [19].

This jump solution still holds for the case for the system with drift

$$\dot{x} = f(x) + g(x)u. \tag{4}$$

The explicit Euler discretization scheme with approximation of $u(t) = u_0 \delta(t)$ by $\tilde{u}(t)$ in $(0, \tau)$ such that $\int_0^\tau \tilde{u}(t)\,\mathrm{d}t = u_0$ is

$$x_{k+1} = x_k + \epsilon [f(x_k) + g(x_k) \tilde{u}_k] \tag{5}$$

with

$$N\epsilon = \tau, \quad \text{and} \quad \epsilon \sum_{k=0}^{N-1} \tilde{u}_k = u_0,$$

and where $x_k \approx x(k\epsilon)$. Note that, since

$$\epsilon \left| \sum_k g(x_k) \tilde{u}_k \right| \le \epsilon \max_{0 \le i \le N} |g(x_i)| \left| \sum_k \tilde{u}_k \right| \le \max_{0 \le i \le N} g(x_i) u_0,$$

it follows that when $\tau \to 0$, the solution of (3) converges to the solution to the drift free equation. Consequently, it still holds that for the general case (4)

$$x(0+) = \sigma_{x(0-)} \mathrm{e}^{u_0 L_g}\, \mathrm{id}, \tag{6}$$

with the same caveat (Frobenius) for extensions to the multi-input case.

In the absence of this Frobenius condition, it is necessary to further discriminate the Dirac delta by using ideas from non-standard analysis. For instance, for $k = 1, 2, \ldots$, let $\delta_{(k)}(t)$ be the weak limit of the sequence $kN^k t^{k-1} \chi_{[0,1/N]}(t)$ for $N \to \infty$ (χ_I is the indicator function of interval I). Clearly, each $\int_{0-}^{0+} \delta_{(k)}(t)\,\mathrm{d}t = 1$, so that their effect is impervious in the linear case. This results in a higher-dimensional characterization for generalized functions (at each singularity time, t_i, an auxiliary function $u_i(\cdot)$, or better a sequence of such functions is specified). In principle, with this extension, powers are Dirac deltas and other multiplicative combinations can formally be specified. For instance $\dot{H} = \delta_{(1)}$ and $H(t)\delta_{(1)}(t) = \frac{1}{2}\delta_{(2)}(t)$, which is consistent with $\frac{\mathrm{d}}{\mathrm{d}t} H^2(t) = 2H(t) \frac{\mathrm{d}}{\mathrm{d}t} H(t)$. Likewise the paradoxical evaluation of $\frac{1}{t} t\delta(t)$ either as $\frac{1}{t}[t\delta(t)] = 0$ or $[\frac{1}{t} t]\delta(t) = \delta(t)$ gets resolved: $t\delta_{(k)}(t) = \lim_{N \to \infty} \frac{k}{(k+1)N} \delta_{(k+1)}$.

V. DERIVATIVES OF IMPULSES

One may wonder if a similar limit approach could be used to model the effect of higher order impulsive inputs (i.e., derivatives of impulses as input). We show in this section that the usual limit procedure, which works in the linear case, must break down, as the associated Lie series does not converge in general. Consider the case of $u = u_1 \dot{\delta}$, where $u_1 \in \mathbb{R}$. Representing the doublet as the limit $\dot{\delta}(t) \approx [\delta(t) -$

$\delta(t-\epsilon)]/\epsilon$, its effect on the state at $\epsilon(+)$ is, in view of the solution for the impulsive input we obtained

$$x(\epsilon+) = \sigma_{x(0-)}\mathrm{e}^{-\frac{u_1}{\epsilon}L_g}\mathrm{e}^{\epsilon L_f}\mathrm{e}^{\frac{u_1}{\epsilon}L_g}\mathrm{id}.$$

The first few terms in the series expansion are

$$(1-\frac{u_1}{\epsilon}+\ldots)(1+\epsilon L_f+\ldots)(1-\frac{u_1}{\epsilon}+\ldots)$$

which give

$$1+\epsilon L_f+(L_fL_g-L_gL_f)-\frac{u_1^2}{\epsilon^2}L_g^2+\ldots$$

i.e., a series results having both positive and negative powers of ϵ, unless the vector field g is independent of the state x. Hence any finite term based computation will lead to erroneous results. One exception is the case where the vector field g is *uniform* over the state space: i.e., when its Jacobian vanishes everywhere. In this case the system is of the form

$$\dot{x}(t) = f(t,x(t)) + g(t)u(t)$$

and L_g is simply a time-variant (or constant as the case may be) vector. Then the Lie-series has only two nonvanishing terms

$$\sigma_x\mathrm{e}^{u_0L_g}\mathrm{id} = x + u_0g,$$

and the effect of the input $[u_1\delta(t)-u_1\delta(t-\epsilon)]/\epsilon$, approximating the doublet is

$$\begin{aligned}
x(\epsilon+) &= \sigma_{x(0-)}\sigma_{x(0-)}\mathrm{e}^{-\frac{u_1}{\epsilon}L_g}\mathrm{e}^{\epsilon L_f}\mathrm{e}^{\frac{u_1}{\epsilon}L_g}\mathrm{id}\\
&= \left(1-\frac{u_1}{\epsilon}L_g\right)\left(1+\epsilon L_f+\frac{\epsilon^2}{2!}L_f^2+\ldots\right)\\
&\quad\times\left(1+\frac{u_1}{\epsilon}L_g\right)\mathrm{id}\\
&= \mathrm{id}+u_1(L_fL_g-L_gL_f)\,\mathrm{id}+h.o.t.
\end{aligned}$$

where we used the fact that $L_fL_g=0$ and $L_g^2=0$, and $h.o.t.$ are the higher order terms. Finally, note that

$$L_fL_g - L_gL_f = L_{[f,g]}$$

where $[f,g]$ is the standard Lie-bracket of the vector fields f and g. But, since g does not depend on the state,

$$[f,g] = J_f(x)g,$$

where J_f is the Jacobian, $\frac{\partial f}{\partial x}$ of f. Unless this Jacobian is proportional to the identity matrix, a new direction in which the state may change is created. The effect of a doublet $u_1\dot{\delta}$ is then obtained by letting $\epsilon\to 0$, to get

$$x(0+) = x(0-) + u_1J_f(x(0-))g. \qquad (7)$$

Likewise the effect of the higher order derivatives of a Dirac delta is

$$u(t) = u_k\delta^{(k)}(t) \implies x(0+) = x(0-) + \sigma_{x(0-)}L_{\mathrm{ad}_f^kg}\,\mathrm{id},$$

where $\mathrm{ad}_f^{k+1}=[f,\mathrm{ad}_f^k]$ and $\mathrm{ad}_fg=[f,g]$, This generalizes results in [6], [7], [8], [9], [13] and may be of interest in optimal feedback control problems as discussed in these papers.

VI. EXAMPLE

Consider the nonlinear model of the pendulum, with a force control whose interaction depends on the excursion of the pendulum. Thus

$$\begin{aligned}
\dot{\theta} &= \omega\\
\dot{\omega} &= -\sin\theta + \beta(\theta)u
\end{aligned}$$

Note that in this case

$$L_gh = \begin{bmatrix} \frac{\partial h_1}{\partial\theta} & \frac{\partial h_1}{\partial\omega}\\ \frac{\partial h_2}{\partial\theta} & \frac{\partial h_2}{\partial\omega}\end{bmatrix}\begin{bmatrix}0\\ \beta(\theta)\end{bmatrix}$$

The effect of an impulsive input $u_0\delta(t)$ is

$$\begin{bmatrix}\theta(0+)\\ \omega(0+)\end{bmatrix} = \sigma_{x(0-)}\mathrm{e}^{u_0L_g}\,\mathrm{id} = \begin{bmatrix}\theta(0-)\\ \omega(0-)\end{bmatrix} + u_0\begin{bmatrix}0\\ \beta(\theta)\end{bmatrix}.$$

Higher order terms in the series vanish. The state can not instantaneously be changed arbitrarily using this impulsive control, but it is clear that if β is not zero at θ_0, then the position at any arbitrary *later* time can be made arbitrary since the initial speed can be changed, thus changing the kinetic energy of the pendulum.
Let's look at the problem now in another way: Assume that β is an invertible function of θ with inverse $\beta = b$, such that $\xi=\beta(\theta)\Rightarrow\theta=b(\xi)$. Changing coordinates, we get

$$\begin{aligned}
\dot{\xi} &= \beta'(b(\xi))\omega\\
\dot{\omega} &= -\sin(b(\xi)) + \xi u.
\end{aligned}$$

The effect of the impulsive input $u=u_0\delta$ is

$$\begin{bmatrix}\xi(0+)\\ \omega(0+)\end{bmatrix} = \begin{bmatrix}\xi(0-)\\ \omega(0-)\end{bmatrix} + u_0\begin{bmatrix}0\\ \xi(0-)\end{bmatrix}.$$

Let now β be independent of θ. In this case we can compute the effect of the control $u=u_0\delta+u_1\dot{\delta}$ by taking first a purely impulsive approximation

$$\tilde{u}_\epsilon(t) = \left(u_0+\frac{u_1}{\epsilon}\right)\delta(t) - \frac{u_1}{\epsilon}\delta(t-\epsilon).$$

Next, let $\epsilon\to 0$ on he system in the original coordinates. Its effect on the state is easily computed

$$\begin{aligned}
\begin{bmatrix}\theta(\epsilon+)\\ \omega(\epsilon+)\end{bmatrix} &= \sigma_{x(0-)}\mathrm{e}^{-\frac{u_1}{\epsilon}L_g}\mathrm{e}^{\epsilon L_f}\mathrm{e}^{(u_0+\frac{u_1}{\epsilon})L_g}\,\mathrm{id}\\
&\to \begin{bmatrix}\theta(0-)-u_1\\ \omega(0-)+u_0\end{bmatrix}.
\end{aligned}$$

VII. GRAMIANS

It was shown that the strength of the impulsive input, u_0, has the same effect on the system $\dot{x}=f(x)+g(x)u$ as when an auxiliary autonomous system $\dot{\xi}=g(x)$ were propagated for a duration u_0. This suggests using an equivalent *flow* for the impulse driven system. This not only justifies the usual description of impulsive systems via the *jump*-maps $x(0+)=F(x(0-),u_0)$, but also allows a more causal interpretation of them. Alternatively, we obtained the ODE associated with impulsive inputs at $t=0$

$$\frac{\partial x(0+,u)}{\partial u} = g(x(0+),u).$$

Suppose we observe $h(x)$, then the *energy* in the transition due to an impulse of strength U may be represented by the quadratic form $\int_0^U h(\mathrm{e}^{\nu L_g}\,\mathrm{id})^2 \mathrm{d}\,\nu$.

This leads one to quantify the degree of instantaneous controllability from the initial state x by a *nonlinear* Gramian

$$\int_0^{u_0} (\mathrm{e}^{\nu L_g} x)(\mathrm{e}^{\nu L_g} x)^\top \,\mathrm{d}\nu$$

which, unlike the linear case, is in general not quadratic in x, due to the differential action of L_g. For instance, for the system

$$\dot{x} = f(x) + Bxu$$

the Lie-exponential is (see Appendix B)

$$\mathrm{e}^{\nu L_g} x = \mathrm{e}^{u_0 B} x,$$

so that the corresponding Gramian is given by

$$P(x) = \int_0^U \mathrm{e}^{B\nu} x x^\top \mathrm{e}^{B^\top \nu}\,\mathrm{d}\nu.$$

Note that strictly speaking, this is an *observability* Gramian which satisfies the Lyapunov equation

$$BP(x) + P(x)B^\top + xx^\top = \mathrm{e}^{BU} x x^\top \mathrm{e}^{B^\top U}.$$

However, it is not clear what role (if any) this Gramians plays or why it might be useful. One possibility may be their relevance for stability analysis [4]. Another avenue to explore is the extension to infinite dimensional systems, particularly systems with delay. It is expected that functional analytic methods such as in [10], [26] will play an important role.

APPENDIX

A. Fine Structure

In this section, we provide a finer structure which discerns between seemingly equivalent representations of generalized functions. By enabling this difference at a microscopic level, the paradoxes stemming from the basic distribution theory can be eliminated. The methodology used is based on generating larger sets from small sets, by considering sequences of elements of the original set. This however may be too rich, and may lead to other problems. Hence the next step is to introduce an equivalence relation in this set of sequences, allowing to group equivalent sequences into equivalence classes. Next we extend the algebraic operations to the equivalence classes, and show that an order relation can be defined on this set of equivalence classes. The original set is embedded in this set of equivalence classes. If the set of equivalence classes of sequences is larger than the original set, then we have extended the class of objects that may be considered. Such a construction is for instance a standard way to define the real numbers as (Cauchy) sequences of rational numbers. We will utilize this first to define the non-standard reals from the reals, the basic construction in non-standard analysis [16], [22].

1) Non-standard reals: Consider the set $\mathbb{R}^{\mathbb{N}}$ of sequences of reals. Define the algebraic operations of addition and multiplication via

$$\begin{aligned} \langle x_n \rangle + \langle y_n \rangle &= \langle x_n + y_n \rangle \\ \langle x_n \rangle \cdot \langle y_n \rangle &= \langle x_n y_n \rangle. \end{aligned}$$

First let us define an additive $\{0,1\}$-valued measure on the set of sub sets of $\mathbb{N}$.

$$\begin{aligned} &m(\mathbb{N}) = 1 \\ &m(A) = 0 \quad \text{if } A \text{ is finite.} \end{aligned}$$

Define an equivalence relation $\sim$ on $\mathbb{R}^{\mathbb{N}}$:

$$\langle x_n \rangle \sim \langle y_n \rangle \Leftrightarrow m\{n | x_n \neq y_n\} = 0.$$

The set $\mathbb{R}^{\mathbb{N}}/_\sim$ is known as the set of non-standard reals and is denoted by ${}^*\mathbb{R}$ [16]. The (standard) reals are embedded in ${}^*\mathbb{R}$ by the injection $\mathbb{R} \hookrightarrow {}^*\mathbb{R} : r \to \langle r, r, r, \ldots \rangle$. With the algebraic operations inherited from the ones defined on $\mathbb{R}^{\mathbb{N}}$, it follows that there are no zero-divisors.
Proof: Let $\langle x_n \rangle \langle y_n \rangle \sim \langle 0 \rangle$. This implies $m\{n | x_n y_n = 0\} = 1$. But

$$\{n | x_n y_n = 0\} = \{n | x_n = 0\} \cup \{n | y_n = 0\}$$

so that $1 \leq m\{n | x_n = 0\} + m\{n | y_n = 0\}$. Since m is a $\{0,1\}$-valued measure, this implies that either $\langle x_n \rangle$ or $\langle y_n \rangle$ is zero. $\square$
The set ${}^*\mathbb{R}$ is ordered by

$$\langle x_n \rangle < \langle y_n \rangle \Leftrightarrow m\{n | x_n < y_n\} = 1.$$

2) *Generalized functions: First we define the *scaling operators*, $\mathbf{S}_\alpha : \mathbb{R}^{\mathbb{R}} \to \mathbb{R}^{\mathbb{R}}$, defined for all $\alpha > 0$ by

$$\sigma_t \mathbf{S}_\alpha x = \sigma_{\alpha t} x = x(\alpha t).$$

Consider a set of smooth functions with support $(0,1)$. Construct from such f the sequence of functions $\langle f_n \rangle$, where $f_n = \mathbf{S}_{\alpha_n} f$ and $\langle \alpha_n \rangle$ is an increasing unbounded sequence of integers (e.g. $\alpha_n = n$ or $\alpha_n = 2^n$). The equivalence class of such a sequence of functions will be defined as a *generalized function and denoted by ${}^*f$. If g is a piecewise continuous (regular) function with a discontinuity at $t = 0$, choose f such that $f(0) = g(0-)$ and $f(1) = g(0+)$, and consider the *generalized function $g +^* f$, which is represented in a small interval $(-\epsilon, \epsilon)$ about 0 by the sequence $\langle g_n \rangle$ with $g_n = g\chi_{(-\epsilon,0)} + f_n \chi_{(0,1/\alpha_n)} + g\chi_{(1/\alpha_n,\epsilon)}$ for sufficiently large n. Here χ_I is the indicator function for the interval I. The main idea behind this construction is that g is represented by a sequence of continuous functions. More generally, if f is piecewise smooth, it can be represented by a sequence of smooth functions. Seen somewhat differently, we can say that a raccordation between the smooth parts is made by introducing a continuation through some "insensible" time. The reason we coined the notion of (in)sensibility is because the evaluation of a *generalized function is defined as follows. For any $t \in *\mathbb{R}$, define the evaluation functional $\sigma_t = \sigma_{\langle t_n \rangle}$ on *generalized functions by

$$\sigma_{\langle t_n \rangle} \langle f_n \rangle = \langle \sigma_{t_n} f_n \rangle = \langle f_n(t_n) \rangle \in {}^*\mathbb{R}.$$

Times that are real ($t \in \mathbb{R}$) are called *sensible*, whereas the argument t of f in $\langle 1/\alpha, f\rangle$ for $\alpha > 1$ is called *insensible* since effectively this *generalized function has no interval as support. However, this insensible time really exists, as illustrated with the following fact: Consider for $0 < t < 1$ and $\alpha > 1$ the Krylov sequence $\langle t_n\rangle = \langle t/\alpha^n\rangle = \langle 1/\alpha; t\rangle$, which represents an *infinitesimal* in the terminology of non-standard analysis, and the Krylov sequence $\langle \alpha^m \mathbf{S}_\alpha; f\rangle$ which represents a *generalized function. Then

$$\begin{aligned} \sigma_{\langle 1/\alpha;t\rangle}\langle \alpha^m \mathbf{S}_\alpha; f\rangle &= \sigma_{\langle t/\alpha^n\rangle}\langle \alpha^{nm}\mathbf{S}_{\alpha^n} f\rangle \\ &= \langle \sigma_{\langle t/\alpha^n}\alpha^{nm}\mathbf{S}_{\alpha^n} f\rangle \\ &= \langle \alpha^{nm} f(t/\alpha^n)\rangle. \end{aligned}$$

If $f(t) = t^k$, the above evaluation gives $\langle \alpha^{nm}t^k/\alpha^{nk}\rangle = t^k\langle \alpha^{m-k}\rangle$. This is either an infinitessimal (for $m < k$), t^k if $m = k$, or an infinite non-standard number if $m > k$. We conclude that a fine-structure evaluation of the *generalized function is possible. The *sensible* times are the $\mathbb{R}$-valued times insofar they are not times of discontinuities and/or singularities. Analogously, we consider the f in the Krylov sequence $\langle \alpha^m \mathbf{S}_\alpha; f\rangle$ the *insensible* part of the *generalized function. It represents the fine-structure of the representation. Moreover, if g_1 is piecewise differentiable (with left and right derivatives defined at 0 and 1), then we define the derivative as

$$\mathbf{D}\langle g_n\rangle = \langle \mathbf{D}g_n\rangle.$$

This implies

$$\mathbf{D}\langle g_n\rangle = \langle \mathbf{D}g_n\rangle = \langle \mathbf{D}\mathbf{S}_{\alpha_n} g_1\rangle = \langle \alpha_n \mathbf{S}_{\alpha_n}\mathbf{D}g_1\rangle.$$

Although the general case above is defined, limiting the α sequences to geometric sequences has some notational benefits, since then $\mathbf{S}_{\alpha_n} = \mathbf{S}_{\alpha^n} = \mathbf{S}_\alpha^n$, and the sequence $\langle \mathbf{S}_{\alpha_n} f\rangle$ is entirely specified by the scaling operator $\mathbf{S}_\alpha$ and the smooth function f, i.e., the Krylov sequence $\langle \mathbf{S}_\alpha; f\rangle$. At once we can then introduce other generalized functions such as $\langle \alpha^m \mathbf{S}_\alpha; f\rangle$ for any $m \in \mathbb{N}$, and note that

$$\begin{aligned} \mathbf{D}\langle \alpha^m \mathbf{S}_\alpha; f\rangle &= \langle \alpha^{m+1}\mathbf{S}_\alpha; \mathbf{D}f\rangle \\ \langle \alpha^p \mathbf{S}_\alpha; g\rangle \cdot \langle \alpha^q \mathbf{S}_\alpha; h\rangle &= \langle \alpha^{p+q}\mathbf{S}_\alpha; gh\rangle. \end{aligned}$$

With this construction the paradox of the Heaviside functions disappears, as H^k and H are clearly different *generalized functions, their insensible parts respectively being $t^k\chi_{(0,1)}$ and $t\chi_{(0,1)}$. Likewise, their derivatives amount to different impulse representations $\delta_{(k)} = \mathbf{D}H^k = kH^{k-1}\delta_{(1)}$, discernible only by their fine structure.

Let us now revisit the *generalized differential equation

$$\mathbf{D}x = xu_0\delta_{(k)}, \quad x(0-) = x_0,$$

We show two methods to prove that for $t > 0$, $x(t) = \mathrm{e}^{u_0}x(0-)$ regardless of k, i.e., the representation of the impulse.

First, search for a series solution of the form $x = \sum_{i=0}^\infty x_i H^i$. One easily deduces that

$$\mathbf{D}x = \sum_{i=1}^\infty x_i H^{i-1}\delta_{(1)} = \sum_{i=0}^\infty x_{i+1}H^i\delta_{(1)}.$$

Likewise,

$$\begin{aligned} u_0 x\delta_{(k)} &= u_0\sum_{i=0}^\infty x_i H^i kH^{k-1}\delta_{(1)} \\ &= u_0 k\sum_{i=k-1}^\infty x_{i-k+1}H^i\delta_{(1)} \end{aligned}$$

Substituting in the *generalized ODE:

$$\begin{aligned} &\sum_{i=0}^{k-2}(i+1)x_{i+1}H^i\delta_{(1)}+ \\ &\quad + \sum_{i=k-1}^\infty [(i+1)x_{i+1} - u_0 kx_{i-k+1}]H^i\delta_{(1)} = 0 \end{aligned}$$

and exploiting the linear independence of the $H^i\delta_{(1)}$, we deduce $x_i = 0$ for $i = 1, 2, \ldots, k-1$, and for $i \geq k$, $x_i = \frac{u_0 k}{i}x_{i-k}$. Note that from the initial condition, $x(0) = x(0-)$. We conclude that $x_{mk} = \frac{u_0^m}{m!}x_0$, and $x_i = 0$ if i is not an m-tuple. But then

$$x = \sum_{m=0}^\infty \frac{u_0^m}{m!}x_0 H^m k = x_0\mathrm{e}^{u_0 H^k}.$$

For a sensible time $t \in \mathbb{R}\setminus\{0\}$, this gives

$$x(t) = \sigma_t x_0\mathrm{e}^{u_0 H} = \begin{cases} x_0 & \text{if} \quad t < 0 \\ \mathrm{e}^{u_0}x_0 & \text{if} \quad t > 0. \end{cases}$$

□

A second shorter but equivalent proof is as follows: Fixing $\alpha > 1$, consider $x = \langle S_\alpha; x_1\rangle$. Then

$$\mathbf{D}x = \langle \alpha S_\alpha; \mathbf{D}x_1\rangle.$$

Likewise,

$$\begin{aligned} u_0 x\delta_{(k)} &= u_0\langle S_\alpha; x_1\rangle k\langle \mathbf{S}_\alpha; H_1^{k-1}\rangle\langle \alpha\mathbf{S}_\alpha; \mathbf{D}H_1\rangle \\ &= u_0 k\langle \alpha\mathbf{S}_\alpha; x_1 H_1^{k-1}\mathbf{D}H_1\rangle \\ &= u_0\langle \alpha\mathbf{S}_\alpha; x_1\mathbf{D}H_1^k\rangle. \end{aligned}$$

Hence

$$\mathbf{D}x_1 = u_0 x_1\mathbf{D}H_1^k,$$

from which for $0 < t < 1$, $\mathbf{D}\log x_1 = u_0\mathbf{D}H_1^k$, and thus

$$x_1(t) = x_1(0-)\mathrm{e}^{u_0 t^k}.$$

Thus, in terms of the sensible time, $0+$: $x(0+) = x_1(1) = x_1(0-)\mathrm{e}^{u_0} = x(0-)\mathrm{e}^{u_0}$, corroborating what we already established by separation of variables.

Finally, the Frobenius condition in the multivariable case implies that all solutions will also be equal for all representations of the impulses, but in the absence of this Frobenius condition we still find a *unique* solution $x_{k_1,\ldots,k_m}$ for each $\delta_{k_1,\ldots,k_m}$.

B. Representation of solutions to nonlinear systems

Using diagrammatic representation, in order to to shed light on the combinatorics, the following was proven in [28].
Theorem: *Given a smooth nonlinear system*

$$\begin{aligned} \dot{x} &= f(x), \quad x(0) = x_o \\ y &= h(x) \end{aligned}$$

where f satisfies a Lipshitz condition, then there exists a T such that for all $t \in [0, T]$ the solution is given by

$$\begin{aligned} x(t) &= \sigma_{x_o} \mathrm{e}^{tL_f} \mathrm{id} \\ y(t) &= \sigma_{x_o} \mathrm{e}^{tL_f} h. \end{aligned}$$

ACKNOWLEDGMENT

The author thanks Nak-seung Patrick Hyun (PhD student) of Georgia Tech for pointing out the relevant references [19], [17], and for valuable discussions during the revision of this paper.

REFERENCES

[1] D.D. Bainov and P.S. Simeonov, *Systems with impulsive effects: Stability, theory and applications*. Academic Press, 1989.
[2] S. Blanes, F. Casas, J.O. Oteo and J. Ros, The Magnus expansion and some of its applications, Physics Reports, 470, 2009, pp. 151–238.
[3] D.M. Bortz and P.W. Nelson, Sensitivity Analysis of a Nonlinear Lumped Parameter Model of HIV Infection Dynamics, Bulletin of Mathematical Biology Vol. 66, (2004) pp. 1009–1026.
[4] C. Briat and A. Seuret, Robust stability of impulsive systems: A functional-based approach, Proc. 4-th IFAC conference on Analysis and Design of Hybrid Systems, Eindhoven, NL, (2012).
[5] C. Briat and E.I. Verriest, A new-delay-SIR model for pulse vaccination. Biomedical Signal Processing and Control, Vol. 4, No. 4, (2009) pp. 272-277.
[6] A.N. Daryin and A.B. Kurzhanski, Impulse Control Inputs and the Theory of Fast Feedback Control, Proc. of 17th World Congress IFAC. 6–11 July 2008. Seoul, 2008. pp. 4869–4874.
[7] A.N. Daryin and A.B. Kurzhanski, Nonlinear feedback types in impulse and fast control, Proc. IFAC NOLCOS, Toulouse, F. 2013.
[8] A.N. Daryin and A.B. Kurzhanski, Control synthesis in a class of higher-order distributions, Differential Equations, Vol. 43, No. 11, 2007, pp. 1479–1489.
[9] A.N. Daryin and Y. Minaeva, Fast controls and their calculation, Proc. of the 19th International Symposium on Mathematical Theory of Networks and Systems, Budapest, H, 2010.
[10] S. Descombes and M. Thalhammer, The Lie-Trotter splitting method for nonlinear evolutionary problems involving critical parameters. An exact local error representation and application to nonlinear Schrödinger equations in the semi-classical regime. (2010) hal-00557593.
[11] I. Halperin, *Introduction to the Theory of Distributions*. University of Toronto Press, 1952.
[12] M. Kuczma, B. Choczewski and R. Ger, *Iterative Functional Equations*. Encyclopedia of Mathematics and its Applications 32, Cambridge University Press, 1990.
[13] A.B. Kurzhanski, On synthesizing impulse controls and the theory of fast controls, Proceedings of the Steklov Institute of Mathematics. 2010. Vol 268 No 1. pp. 207–221. (translated from Russian).
[14] A.B. Kurzhanski and Y.S. Osipov, On controlling linear systems through generalized controls, Differential Equations, Vol. 5, No. 8, 1969, 1360–1370.
[15] D.F. Lawdon, *Optimal Trajectories for Space Navigation*, London, Butterworth, 1963.
[16] T/ Lindstrom, Nonstandard analysis and perturbations of the Laplacian along Brownian motion, in *Stochastic Processes - Mathematics and Physics*, S. Albeverio, Ph. Blancahrd, L. Streit (eds.), Springer-Verlag, Lecture Notes in Mathematics, no. 1158, 1984, pp. 180–200.
[17] B. Miller, The generalized solutions of nonlinear optimization problems with impulse control, SIAM J. Contr. Optimiz. Vol. 34, No. 4, 1996, 1420–1440.
[18] L. Neustadt, A general theory of minimum-fuel trasjectories, SIAM J. Contr. Ser A. Vol. 3, No. 2, 1965, 317–356.
[19] Y. Orlov, Instantaneous impulse response of nonlinear system, IEEE Transactions on Automatic Control, Vol. AC-45, No. 5, 2000, pp 999–1001.
[20] M. Pototschnig, J. Niegemann, L. Tkeshelashvili and K. Busch, Time-domain simulations of the nonlinear Maxwell equations using operator-exponential methods, IEEE Transactions on Antennas and Propagation, Vol. 57, No. 2 (2009), pp. 475.
[21] R.W. Rishel, An extended Pontryagin principle for control systems, whose control laws contain measures, SIAM J. Contr. Ser A. Vol. 3, No. 2, 1965, 191–205.
[22] A. Robinson, *Non-Standard Analysis*, Princeton University Press, 1966.
[23] C. Schwartz, Nonlinear operators and their propagators, J. Math, Phys, 38, 1 1997, pp. 484–500.
[24] C. Schwartz, Nonlinear operators II, J. Math. Phys. 38, 7, 1997, pp. 3841–3862.
[25] L.M. Silverman, Representation and realization of time-variable linear systems, Technical Report No. 94, Department of Electrical Engineering, Columbia University, 1966.
[26] S. Sternberg, Local Contractions and a theorem of Poincare, American Journal of Mathematics, Vol. 79, N0.4 (1957) pp. 809–824.
[27] E.I. Verriest, Regularization method for optimally switched and impulsive systems with biomedical applications, Proceedings of the 42th IEEE Conference on Decision and Control, pp. 2156-2161, Maui, HI, December 2003.
[28] E.I. Verriest, Lie exponentials, tadpole diagrams and frogspawn: explicit solutions to autonomous nonlinear systems, in Dynamic Systems and Applications, Dynamic Publishers, Inc., 5 (2008) pp. 481-488.
[29] E.I. Verriest and T. Kailath, On generalized balanced realizations, IEEE Transactions on Automatic Control, Vol. AC-28, No. 8, 1983, pp 833–844.
[30] E.I. Verriest and P. Pepe, Time optimal and optimal impulsive control for coupled differential difference point delay systems with an application in forestry. in Topics in Time Delay Systems: Analysis, Algorithms and Control, J.-J. Loiseau, W. Michiels, S.I. Niculescu and R. Sipahi, (eds.), Springer-Verlag, Lecture Notes in Control and Information Sciences, Vol. 388, pp. 255-265, 2009.
[31] T. Yang, *Impulsive Control Theory*, Springer-Verlag, 2001.
[32] S.T. Zavalishchin and A.N. Sesekin, *Dynamic Impulse Systems: Theory and Applications*, Kluwer 1997.